



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>6</sup> : <b>C12N 15/12, C07K 14/47, C12N 15/11, C12Q 1/68, C07K 16/18, G01N 33/68, A61K 38/17</b>		A2	(11) International Publication Number: <b>WO 00/00610</b>
			(43) International Publication Date: 6 January 2000 (06.01.00)
(21) International Application Number: PCT/US99/14484		(72) Inventors; and	
(22) International Filing Date: 25 June 1999 (25.06.99)		(75) Inventors/Applicants (for US only): LAL, Preeti [IN/US]; 2382 Lass Drive, Santa Clara, CA 95054 (US). TANG, Y., Tom [CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). GORGONE, Gina, A. [US/US]; 1253 Pinccrest Drive, Boulder Creek, CA 95006 (US). CORLEY, Neil, C. [US/US]; 1240 Dale Avenue #30, Mountain View, CA 94040 (US). GUEGLER, Karl, J. [CH/US]; 1048 Oakland Avenue, Menlo Park, CA 94025 (US). BAUGHN, Mariah, R. [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). AKERBLOM, Ingrid, E. [US/US]; 1234 Johnson Street, Redwood City, CA 94061 (US). AU-YOUNG, Janice [US/US]; 1419 Kains Avenue, Berkeley, CA 94702 (US). YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). PATTERSON, Chandra [US/US]; 490 Sherwood Way #1, Menlo Park, CA 94025 (US). REDDY, Roopa [IN/US]; 1233 W. McKinley Drive, Sunnyvale, CA 94086 (US). HILLMAN, Jennifer, L. [US/US]; 230 Monroe Drive #12, Mountain View, CA 94040 (US). BANDMAN, Olga [US/US]; 366 Anna Avenue, Mountain View, CA 94043 (US).	
(30) Priority Data:			
60/090,762	26 June 1998 (26.06.98)	US	
60/094,983	31 July 1998 (31.07.98)	US	
60/102,686	1 October 1998 (01.10.98)	US	
60/112,129	11 December 1998 (11.12.98)	US	
(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Applications			
US	60/090,762 (CIP)		
Filed on	26 June 1998 (26.06.98)		
US	60/094,983 (CIP)		
Filed on	31 July 1998 (31.07.98)		
US	60/102,686 (CIP)		
Filed on	1 October 1998 (01.10.98)		
US	60/112,129 (CIP)		
Filed on	11 December 1998 (11.12.98)		
(71) Applicant (for all designated States except US): INCYTE PHARMACEUTICALS, INC. [US/US]; 3174 Porter Drive, Palo Alto, CA 94304 (US).		(74) Agents: BILLINGS, Lucy, J. et al.; Incyte Pharmaceuticals, Inc., 3174 Porter Drive, Palo Alto, CA 94304 (US).	
		(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).	
		Published Without international search report and to be republished upon receipt of that report.	
(54) Title: HUMAN SIGNAL PEPTIDE-CONTAINING PROTEINS			
(57) Abstract			
<p>The invention provides human signal peptide-containing proteins (HSPP) and polynucleotides which indentify and encode HSPP. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with expression of HSPP.</p>			

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Licchtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

## HUMAN SIGNAL PEPTIDE-CONTAINING PROTEINS

5

### TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of human signal peptide-containing proteins and to the use of these sequences in the diagnosis, treatment, and prevention of cell proliferative disorders including cancer; inflammation; and cardiovascular, neurological, reproductive, and developmental disorders.

### BACKGROUND OF THE INVENTION

Protein transport is essential for cellular function. Transport of a protein may be mediated by a signal peptide located at the amino terminus of the protein itself. The signal peptide is comprised of about ten to twenty hydrophobic amino acids which target the nascent protein from the ribosome to a particular membrane bound compartment such as the endoplasmic reticulum (ER). Proteins targeted to the ER may either proceed through the secretory pathway or remain in any of the secretory organelles such as the ER, Golgi apparatus, or lysosomes. Proteins that transit through the secretory pathway are either secreted into the extracellular space or retained in the plasma membrane. Secreted proteins are often synthesized as inactive precursors that are activated by post-translational processing events during transit through the secretory pathway. Such events include glycosylation, phosphorylation, proteolysis, and removal of the signal peptide by a signal peptidase. Other events that may occur during protein transport include chaperone-dependent unfolding and folding of the nascent protein and interaction of the protein with a receptor or pore complex. Examples of secreted proteins with amino terminal signal peptides are discussed below and include receptors, extracellular matrix molecules, cytokines, hormones, growth and differentiation factors, neuropeptides, vasomediators, phosphokinases, phosphatases, phospholipases, phosphodiesterases, G and Ras-related proteins, ion channels, transporters/pumps, proteases, and transcription factors. (Reviewed in Alberts, B. et al. (1994) Molecular Biology of The Cell, Garland Publishing, New York, NY, pp. 557-560, 582-592.)

G-protein coupled receptors (GPCRs) comprise a superfamily of integral membrane proteins which transduce extracellular signals. GPCRs include receptors for biogenic amines such as dopamine, epinephrine, histamine, glutamate (metabotropic effect), acetylcholine (muscarinic effect), and serotonin; for lipid mediators of inflammation such as prostaglandins, platelet activating factor, and leukotrienes; for peptide hormones such as calcitonin, C5a anaphylatoxin, follicle stimulating hormone, gonadotropin releasing hormone, neurokinin, oxytocin, and thrombin; and for sensory signal mediators such as retinal photopigments and olfactory stimulatory molecules. The structure of these highly conserved receptors consists of seven hydrophobic transmembrane regions, cysteine disulfide bridges between the second and third extracellular loops, an extracellular N-terminus, and a cytoplasmic C-terminus. The N-terminus interacts with ligands, the disulfide bridges interact with agonists and antagonists, and the large third intracellular loop interacts with G proteins to activate second messengers such as cyclic AMP, phospholipase C, inositol triphosphate, or ion channels. (Reviewed in Watson, S. and Arkinstall, S. (1994) The G-protein Linked Receptor Facts Book, Academic Press, San Diego, CA, pp. 2-6; and Bolander, F.F. (1994) Molecular Endocrinology, Academic Press, San Diego, CA, pp. 162-176.)

Other types of receptors include cell surface antigens identified on leukocytic cells of the immune system. These antigens have been identified using systematic, monoclonal antibody (mAb)-based "shot gun" techniques. These techniques have resulted in the production of hundreds of mAbs directed against unknown cell surface leukocytic antigens. These antigens have been grouped into "clusters of differentiation" based on common immunocytochemical localization patterns in various differentiated and undifferentiated leukocytic cell types. Antigens in a given cluster are presumed to identify a single cell surface protein and are assigned a "CD" number. Some of the genes encoding proteins identified by CD antigens have been isolated and characterized as both transmembrane proteins and cell surface proteins anchored to the plasma membrane via covalent attachment to fatty acid-containing glycolipids such as glycosylphosphatidylinositol (GPI). (Reviewed in Barclay, A. N. et al. (1993) The Leucocyte Antigen Facts Book, Academic Press, San Diego, CA, pp. 144-145; Noel, L. S. et al. (1998) *J. Biol. Chem.* 273:3878-3883.)

Tetraspanins are a superfamily of membrane proteins which facilitate the formation

and stability of cell-surface signaling complexes containing lineage-specific proteins, integrins, and other tetraspanins. They are involved in cell activation, proliferation (including cancer), differentiation, adhesion, and motility. These proteins cross the membrane four times, have conserved intracellular – and C-termini and an extracellular, non-conserved hydrophilic domain. Tetraspanins include, e.g., platelet and endothelial cell membrane proteins, leukocyte surface proteins, tissue specific and tumorous antigens, and the retinitis pigmentosa-associated gene peripherin. (Maecker, H.T. et al. (1997) FASEB J. 11:428-442.)

Matrix proteins (MPs) are transmembrane and extracellular proteins which function in formation, growth, remodeling, and maintenance of tissues and as important mediators and regulators of the inflammatory response. The expression and balance of MPs may be perturbed by biochemical changes that result from congenital, epigenetic, or infectious diseases. In addition, MPs affect leukocyte migration, proliferation, differentiation, and activation in the immune response. MPs are frequently characterized by the presence of one or more domains which may include collagen-like domains, EGF-like domains, immunoglobulin-like domains, and fibronectin-like domains. In addition, some MPs are heavily glycosylated. MPs include extracellular proteins such as fibronectin, collagen, and galectin and cell adhesion receptors such as cell adhesion molecules (CAMs), cadherins, and integrins. (Reviewed in Ayad, S. et al. (1994) The Extracellular Matrix Facts Book, Academic Press, San Diego, CA, pp. 2-16; Ruoslahti, E. (1997) *Kidney Int.* 51:1413-1417; Sjaastad, M.D. and Nelson, W.J. (1997) *BioEssays* 19:47-55.)

Lectins are proteins characterized by their ability to bind carbohydrates on cell membranes by means of discrete, modular carbohydrate recognition domains, CRDs. (Kishore, U. et al. (1997) *Matrix Biol.* 15:583-592.) Certain cytokines and membrane-spanning proteins have CRDs which may enhance interactions with extracellular or intracellular ligands, with proteins in secretory pathways, or with molecules in signal transduction pathways. The lipocalin superfamily constitutes a phylogenetically conserved group of more than forty proteins that function by binding to and transporting a variety of physiologically important ligands. (Tanaka, T. et al. (1997) *J. Biol. Chem.* 272:15789-15795; and van't Hof, W. et al. (1997) *J. Biol. Chem.* 272:1837-1841.)

Selectins are a family of calcium ion-dependent lectins expressed on inflamed vascular

endothelium and the surface of some leukocytes. (Rossiter, H. et al. (1997) Mol. Med. Today 3:214-222.)

Protein kinases regulate many different cell proliferation, differentiation, and signaling processes by adding phosphate groups to proteins. Reversible protein phosphorylation is a key strategy for controlling protein functional activity in eukaryotic cells. The high energy phosphate which drives this activation is generally transferred from adenosine triphosphate molecules (ATP) to a particular protein by protein kinases and removed from that protein by protein phosphatases. Phosphorylation occurs in response to extracellular signals, cell cycle checkpoints, and environmental or nutritional stresses. Protein kinases may be roughly divided into two groups; protein tyrosine kinases (PTKs) which phosphorylate tyrosine residues, and serine/threonine kinases (STKs) which phosphorylate serine or threonine residues. A few protein kinases have dual specificity. A majority of kinases contain a similar 250-300 amino acid catalytic domain. (Hardie, G. and Hanks, S. (1995) The Protein Kinase Facts Book, Vol I, pp. 7-47, Academic Press, San Diego, CA.)

Protein phosphatases remove phosphate groups from molecules previously modified by protein kinases thus participating in cell signaling, proliferation, differentiation, contacts, and oncogenesis. Protein phosphorylation is a key strategy used to control protein functional activity in eukaryotic cells. The high energy phosphate is transferred from ATP to a protein by protein kinases and removed by protein phosphatases. There appear to be three, evolutionarily-distinct protein phosphatase gene families: protein phosphatases (PPs); protein tyrosine phosphatases (PTPs); and acid/alkaline phosphatases (APs). PPs dephosphorylate phosphoserine/threonine residues and are an important regulator of many cAMP mediated, hormone responses in cells. PTPs reverse the effects of protein tyrosine kinases and therefore play a significant role in cell cycle and cell signaling processes. Although APs dephosphorylate substrates in vitro, their role in vivo is not well known. (Charbonneau, H. and Tonks, N.K. (1992) Annu. Rev. Cell Biol. 8:463-493.)

Cyclic nucleotides (cAMP and cGMP) function as intracellular second messengers to transduce a variety of extracellular signals, including hormones, light and neurotransmitters. Cyclic nucleotide phosphodiesterases (PDEs) degrade cyclic nucleotides to their corresponding monophosphates, thereby regulating the intracellular

concentrations of cyclic nucleotides and their effects on signal transduction. At least seven families of mammalian PDEs have been identified based on substrate specificity and affinity, sensitivity to cofactors and sensitivity to inhibitory drugs. (Beavo, J.A. (1995) *Physiological Reviews* 75: 725-748.)

5        Phospholipases (PLs) are enzymes that catalyze the removal of fatty acid residues from phosphoglycerides. PLs play an important role in transmembrane signal transduction and are named according to the specific ester bond in phosphoglycerides that is hydrolyzed, i.e., A<sub>1</sub>, A<sub>2</sub>, C or D. PLA<sub>2</sub> cleaves the ester bond at position 2 of the glycerol moiety of membrane phospholipids giving rise to arachidonic acid. Arachidonic acid is  
10 the common precursor to four major classes of eicosanoids, namely prostaglandins, prostacyclins, thromboxanes and leukotrienes. Eicosanoids are signaling molecules involved in the contraction of smooth muscle, platelet aggregation, and pain and inflammatory responses. (Alberts, B. et al. (1994) Molecular Biology of The Cell, Garland Publishing, Inc., New York, NY, pp. 85, 211, 239-240, 642-645.)

15        The nucleotide cyclases, i.e., adenylate and guanylate cyclase, catalyze the synthesis of the cyclic nucleotides, cAMP and cGMP, from ATP and GTP, respectively. They act in concert with phosphodiesterases, which degrade cAMP and cGMP, to regulate the cellular levels of these molecules and their functions. cAMP and cGMP function as intracellular second messengers to transduce a variety of extracellular signals, e.g.,  
20 hormones, and light and neurotransmitters. (Stryer, L. (1988) Biochemistry W.H. Freeman and Co., New York, pp. 975-980, 1029-1035.)

Cytokines are produced in response to cell perturbation. Some cytokines are produced as precursor forms, and some form multimers in order to become active. They are produced in groups and in patterns characteristic of the particular stimulus or disease,  
25 and the members of the group interact with one another and other molecules to produce an overall biological response. Interleukins, neurotrophins, growth factors, interferons, and chemokines are all families of cytokines which work in conjunction with cellular receptors to regulate cell proliferation and differentiation and to affect such activities as leukocyte migration and function, hematopoietic cell proliferation, temperature regulation, acute  
30 response to infections, tissue remodeling, apoptosis, and cell survival. Studies using antibodies or other drugs that modify the activity of a particular cytokine are used to elucidate the roles of individual cytokines in pathology and physiology.

Chemokines, in particular, are small chemoattractant cytokines involved in inflammation, leukocyte proliferation and migration, angiogenesis and angiostasis, regulation of hematopoiesis, HIV infectivity, and stimulation of cytokine secretion. Chemokines generally contain 70-100 amino acids and are subdivided into four  
5 subfamilies based on the presence of conserved cysteine-based motifs. (Callard, R. and Gearing, A. (1994) The Cytokine Facts Book, Academic Press, New York, NY, pp. 181-190, 210-213, 223-227.)

Growth and differentiation factors are secreted proteins which function in intercellular communication. Some factors require oligomerization or association with  
10 MPs for activity. Complex interactions among these factors and their receptors trigger intracellular signal transduction pathways that stimulate or inhibit cell division, cell differentiation, cell signaling, and cell motility. Most growth and differentiation factors act on cells in their local environment (paracrine signaling). There are three broad classes of growth and differentiation factors. The first class includes the large polypeptide growth  
15 factors such as epidermal growth factor, fibroblast growth factor, transforming growth factor, insulin-like growth factor, and platelet-derived growth factor. The second class includes the hematopoietic growth factors such as the colony stimulating factors (CSFs). Hematopoietic growth factors stimulate the proliferation and differentiation of blood cells such as B-lymphocytes, T-lymphocytes, erythrocytes, platelets, eosinophils, basophils,  
20 neutrophils, macrophages, and their stem cell precursors. The third class includes small peptide factors such as bombesin, vasopressin, oxytocin, endothelin, transferrin, angiotensin II, vasoactive intestinal peptide, and bradykinin which function as hormones to regulate cellular functions other than proliferation.

Growth and differentiation factors play critical roles in neoplastic transformation of  
25 cells in vitro and in tumor progression in vivo. Inappropriate expression of growth factors by tumor cells may contribute to vascularization and metastasis of melanotic tumors. During hematopoiesis, growth factor misregulation can result in anemias, leukemias, and lymphomas. Certain growth factors such as interferon are cytotoxic to tumor cells both in vivo and in vitro. Moreover, some growth factors and growth factor receptors are related  
30 both structurally and functionally to oncoproteins. In addition, growth factors affect transcriptional regulation of both proto-oncogenes and oncosuppressor genes. (Reviewed in Pimentel, E. (1994) Handbook of Growth Factors, CRC Press, Ann Arbor, MI, pp. 1-9.)



Proteolytic enzymes or proteases either activate or deactivate proteins by hydrolyzing peptide bonds. Proteases are found in the cytosol, in membrane-bound compartments, and in the extracellular space. The major families are the zinc, serine, cysteine, thiol, and carboxyl proteases.

- 5 Zinc proteases, e.g., carboxypeptidase A, have a zinc ion bound to the active site. These proteases recognize C-terminal residues that contain an aromatic or bulky aliphatic side chain, and hydrolyze the peptide bond adjacent to the C-terminal residues. Serine proteases have an active site serine residue and include digestive enzymes, e.g., trypsin and chymotrypsin, components of the complement and blood-clotting cascades, and
- 10 enzymes that control the degradation and turnover of extracellular matrix (ECM) molecules. Cysteine proteases (e.g. cathepsin) are produced by monocytes, macrophages and other immune cells, and are involved in diverse cellular processes ranging from the processing of precursor proteins to intracellular degradation. Overproduction of these enzymes can cause the tissue destruction associated with rheumatoid arthritis and asthma.
- 15 Thiol proteases, e.g., papain, contain an active site cysteine and are widely distributed within tissues. Carboxyl proteases, e.g., pepsin, are active only under acidic conditions (pH 2 to 3).

- Guanosine triphosphate-binding proteins (G proteins) can be grouped into two major classes: heterotrimeric G proteins and small G proteins. Heterotrimeric G proteins
- 20 interact with GPCRs that respond to hormones, growth factors, neuromodulators, or other signaling molecules. The interaction between GPCR and G protein allows the G protein to exchange GTP for guanosine diphosphate (GDP). This exchange activates the G protein, allowing it to dissociate from the receptor and interact with its cognate second messenger-generating protein, e.g., adenylate cyclase, guanylate cyclase, phospholipase C,
- 25 or ion channels. The hydrolysis of GTP to GDP by the G protein acts as an on-off switch, terminating the action of the G protein and preparing it to interact with another receptor molecule, thus beginning another round of signal transduction.

- The small G proteins consist of single 21-30 kDa polypeptides. They can be classified into five subfamilies: Ras, Rho, Ran, Rab, and ADP-ribosylation factor. These
- 30 proteins regulate cell growth, cell cycle control, protein secretion, and intracellular vesicle interaction. In particular, the Ras proteins are essential in transducing signals from receptor tyrosine kinases to serine/threonine kinases which control cell growth and

differentiation. Mutant Ras proteins, which bind but can not hydrolyze GTP, are permanently activated and cause continuous cell proliferation or cancer. All five subfamilies share common structural features and four conserved motifs. Most of the membrane-bound G proteins require a carboxy terminal isoprenyl group (CAAX), added posttranslationally, for membrane association and biological activity. The G proteins also have a variable effector region, located between motifs I and II, which is characterized as the interaction site for guanine nucleotide exchange factors or GTPase-activating proteins.

Eukaryotic cells are bound by a membrane and subdivided into membrane-bound compartments. Membranes are impermeable to many ions and polar molecules, therefore transport of these molecules is mediated by ion channels, ion pumps, transport proteins, or pumps. Symporters and antiporters regulate cytosolic pH by transporting ions and small molecules, e.g., amino acids, glucose, and drugs, across membranes; symporters transport small molecules and ions in the same direction, and antiporters, in the opposite direction. Transporter superfamilies include facilitative transporters and active ATP binding cassette transporters involved in multiple-drug resistance and the targeting of antigenic peptides to MHC Class I molecules. These transporters bind to a specific ion or other molecule and undergo conformational changes in order to transfer the ion or molecule across a membrane. Transport can occur by a passive, concentration-dependent mechanism or can be linked to an energy source such as ATP hydrolysis or an ion gradient.

Ion channels, ion pumps, and transport proteins mediate the transport of molecules across cellular membranes. Symporters and antiporters regulate cytosolic pH by transporting ions and small molecules such as amino acids, glucose, and drugs. Symporters transport small molecules and ions unidirectionally, and antiporters, bidirectionally. Transporter superfamilies include facilitative transporters and active ATP-binding cassette transporters which are involved in multiple-drug resistance and the targeting of antigenic peptides to MHC Class I molecules. These transporters bind to a specific ion or other molecule and undergo a conformational change in order to transfer the ion or molecule across the membrane. Transport can occur by a passive, concentration-dependent mechanism or can be linked to an energy source such as ATP hydrolysis. (Reviewed in Alberts, B. et al. (1994) Molecular Biology of The Cell, Garland Publishing, New York, NY, pp. 523-546.)

Ion channels are formed by transmembrane proteins which create a lined passageway across the membrane through which water and ions, such as  $\text{Na}^+$ ,  $\text{K}^+$ ,  $\text{Ca}^{2+}$ , and  $\text{Cl}^-$ , enter and exit the cell. For example, chloride channels are involved in the regulation of the membrane electric potential as well as absorption and secretion of ions across the membrane. Chloride channels also regulate the internal pH of membrane-bound organelles.

Ion pumps are ATPases which actively maintain membrane gradients. Ion pumps are classified as P, V, or F according to their structure and function. All have one or more binding sites for ATP in their cytosolic domains. The P-class ion pumps include  $\text{Ca}^{2+}$  ATPase and  $\text{Na}^+/\text{K}^+$  ATPase and function in transporting  $\text{H}^+$ ,  $\text{Na}^+$ ,  $\text{K}^+$ , and  $\text{Ca}^{2+}$  ions. P-class pumps consist of two  $\alpha$  and two  $\beta$  transmembrane subunits. The V- and F-class ion pumps have similar structures and but transport only  $\text{H}^+$ . F class  $\text{H}^+$  pumps mediate transport across the membranes of mitochondria and chloroplasts, while V-class  $\text{H}^+$  pumps regulate acidity inside lysosomes, endosomes, and plant vacuoles.

A family of structurally related intrinsic membrane proteins known as facilitative glucose transporters catalyze the movement of glucose and other selected sugars across the plasma membrane. The proteins in this family contain a highly conserved, large transmembrane domain comprised of 12  $\alpha$ -helices, and several weakly conserved, cytoplasmic and exoplasmic domains (Pessin, J. E., and Bell, G.I. (1992) *Annu. Rev. Physiol.* 54:911-930).

Amino acid transport is mediated by  $\text{Na}^+$  dependent amino acid transporters. These transporters are involved in gastrointestinal and renal uptake of dietary and cellular amino acids and in neuronal reuptake of neurotransmitters. Transport of cationic amino acids is mediated by the system  $\gamma^+$  family and the cationic amino acid transporter (CAT) family. Members of the CAT family share a high degree of sequence homology, and each contains 12-14 putative transmembrane domains (Ito, K. and Groudine, M. (1997) *J. Biol. Chem.* 272:26780-26786).

Proton-coupled, 12 membrane-spanning domain transporters such as PEPT 1 and PEPT 2 are responsible for gastrointestinal absorption and for renal reabsorption of peptides using an electrochemical  $\text{H}^+$  gradient as the driving force. A heterodimeric peptide transporter, consisting of TAP 1 and TAP 2, is associated with antigen processing. Peptide antigens are transported across the membrane of the endoplasmic reticulum so

they can be presented to the major histocompatibility complex class I molecules. Each TAP protein consists of multiple hydrophobic membrane spanning segments and a highly conserved ATP-binding cassette. (Boll, M. et al. (1996) Proc. Natl. Acad. Sci. 93:284-289.)

- 5           Hormones are secreted molecules that travel through the circulation and bind to specific receptors on the surface of, or within, target cells. Although they have diverse biochemical compositions and mechanisms of action, hormones can be grouped into two categories. One category consists of small lipophilic hormones that diffuse through the plasma membrane of target cells, bind to cytosolic or nuclear receptors, and form a
- 10 complex that alters gene expression. Examples of these molecules include retinoic acid, thyroxine, and the cholesterol-derived steroid hormones such as progesterone, estrogen, testosterone, cortisol, and aldosterone. The second category consists of hydrophilic hormones that function by binding to cell surface receptors that transduce signals across the plasma membrane. Examples of such hormones include amino acid derivatives such
- 15 as catecholamines and peptide hormones such as glucagon, insulin, gastrin, secretin, cholecystokinin, adrenocorticotrophic hormone, follicle stimulating hormone, luteinizing hormone, thyroid stimulating hormone, and vasopressin. (See, for example, Lodish et al. (1995) Molecular Cell Biology, Scientific American Books Inc., New York, NY, pp. 856-864.)
- 20           Neuropeptides and vasomediators (NP/VM) comprise a large family of endogenous signaling molecules. Included in this family are neuropeptides and neuropeptide hormones such as bombesin, neuropeptide Y, neurotensin, neuromedin N, melanocortins, opioids, galanin, somatostatin, tachykinins, urotensin II and related peptides involved in smooth muscle stimulation, vasopressin, vasoactive intestinal peptide,
- 25 and circulatory system-borne signaling molecules such as angiotensin, complement, calcitonin, endothelins, formyl-methionyl peptides, glucagon, cholecystokinin and gastrin. NP/VMs can transduce signals directly, modulate the activity or release of other neurotransmitters and hormones, and act as catalytic enzymes in cascades. The effects of NP/VMs range from extremely brief to long-lasting. (Reviewed in Martin, C. R. et al.
- 30 (1985) Endocrine Physiology, Oxford University Press, New York, NY, pp. 57-62.)

Regulatory molecules turn individual genes or groups of genes on and off in response to various inductive mechanisms of the cell or organism; act as transcription factors by determining

whether or not transcription is initiated, enhanced, or repressed; and splice transcripts as dictated in a particular cell or tissue. Although they interact with short stretches of DNA scattered throughout the entire genome, most gene expression is regulated near the site at which transcription starts or within the open reading frame of the gene being expressed. Many of the transcription factors incorporate one of a set of DNA-binding structural motifs, each of which contains either  $\alpha$  helices or  $\beta$  sheets and binds to the major groove of DNA. (Pabo, C.O. and R.T. Sauer (1992) *Ann. Rev. Biochem.* 61:1053-95.) Other domains of transcription factors may form crucial contacts with the DNA. In addition, accessory proteins provide important interactions which may convert a particular protein complex to an activator or a repressor or may prevent binding. (Alberts, B. et al. (1994) Molecular Biology of the Cell, Garland Publishing Co, New York, NY pp. 401-474.)

The discovery of new human signal peptide-containing proteins and the polynucleotides encoding them satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of cell proliferative disorders including cancer; inflammation; and cardiovascular, neurological, reproductive, and developmental disorders.

### SUMMARY OF THE INVENTION

The invention features substantially purified polypeptides, proteins with signal peptides, referred to collectively as "HSPP" and individually as "HSPP-1", "HSPP-2", "HSPP-3", "HSPP-4", "HSPP-5", "HSPP-6", "HSPP-7", "HSPP-8", "HSPP-9", "HSPP-10", "HSPP-11", "HSPP-12", "HSPP-13", "HSPP-14", "HSPP-15", "HSPP-16", "HSPP-17", "HSPP-18", "HSPP-19", "HSPP-20", "HSPP-21", "HSPP-22", "HSPP-23", "HSPP-24", "HSPP-25", "HSPP-26", "HSPP-27", "HSPP-28", "HSPP-29", "HSPP-30", "HSPP-31", "HSPP-32", "HSPP-33", "HSPP-34", "HSPP-35", "HSPP-36", "HSPP-37", "HSPP-38", "HSPP-39", "HSPP-40", "HSPP-41", "HSPP-42", "HSPP-43", "HSPP-44", "HSPP-45", "HSPP-46", "HSPP-47", "HSPP-48", "HSPP-49", "HSPP-50", "HSPP-51", "HSPP-52", "HSPP-53", "HSPP-54", "HSPP-55", "HSPP-56", "HSPP-57", "HSPP-58", "HSPP-59", "HSPP-60", "HSPP-61", "HSPP-62", "HSPP-63", "HSPP-64", "HSPP-65", "HSPP-66", "HSPP-67", "HSPP-68", "HSPP-69", "HSPP-70", "HSPP-71", "HSPP-72", "HSPP-73", "HSPP-74", "HSPP-75", "HSPP-76", "HSPP-77", "HSPP-78", "HSPP-79", "HSPP-80", "HSPP-81", "HSPP-82", "HSPP-83", "HSPP-84", "HSPP-85", "HSPP-86", "HSPP-87", "HSPP-88", "HSPP-89", "HSPP-90", "HSPP-91", "HSPP-92", "HSPP-93", "HSPP-94", "HSPP-95", "HSPP-96", "HSPP-97", "HSPP-98", "HSPP-99", "HSPP-100", "HSPP-

101", "HSPP-102", "HSPP-103", "HSPP-104", "HSPP-105", "HSPP-106", "HSPP-107",  
"HSPP-108", "HSPP-109", "HSPP-110", "HSPP-111", "HSPP-112", "HSPP-113", "HSPP-  
114", "HSPP-115", "HSPP-116", "HSPP-117", "HSPP-118", "HSPP-119", "HSPP-120",  
"HSPP-121", "HSPP-122", "HSPP-123", "HSPP-124", "HSPP-125", "HSPP-126",  
5 "HSPP-127", "HSPP-128", "HSPP-129", "HSPP-130", "HSPP-131", "HSPP-132",  
"HSPP-133", and "HSPP-134". In one aspect, the invention provides a substantially  
purified polypeptide comprising an amino acid sequence selected from the group  
consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5,  
SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID  
10 NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16,  
SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ  
ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID  
NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID  
NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37,  
15 SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:42, SEQ  
ID NO:43, SEQ ID NO:44, SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID  
NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:53,  
SEQ ID NO:54, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:57, SEQ ID NO:58, SEQ  
ID NO:59, SEQ ID NO:60, SEQ ID NO:61, SEQ ID NO:62, SEQ ID NO:63, SEQ ID  
20 NO:64, SEQ ID NO:65, SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69,  
SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74, SEQ  
ID NO:75, SEQ ID NO:76, SEQ ID NO:77, SEQ ID NO:78, SEQ ID NO:79, SEQ ID  
NO:80, SEQ ID NO:81, SEQ ID NO:82, SEQ ID NO:83, SEQ ID NO:84, SEQ ID NO:85,  
SEQ ID NO:86, SEQ ID NO:87, SEQ ID NO:88, SEQ ID NO:89, SEQ ID NO:90, SEQ  
25 ID NO:91, SEQ ID NO:92, SEQ ID NO:93, SEQ ID NO:94, SEQ ID NO:95, SEQ ID  
NO:96, SEQ ID NO:97, SEQ ID NO:98, SEQ ID NO:99, SEQ ID NO:100, SEQ ID  
NO:101, SEQ ID NO:102, SEQ ID NO:103, SEQ ID NO:104, SEQ ID NO:105, SEQ ID  
NO:106, SEQ ID NO:107, SEQ ID NO:108, SEQ ID NO:109, SEQ ID NO:110, SEQ ID  
NO:111, SEQ ID NO:112, SEQ ID NO:113, SEQ ID NO:114, SEQ ID NO:115, SEQ ID  
30 NO:116, SEQ ID NO:117, SEQ ID NO:118, SEQ ID NO:119, SEQ ID NO:120, SEQ ID  
NO:121, SEQ ID NO:122, SEQ ID NO:123, SEQ ID NO:124, SEQ ID NO:125, SEQ ID  
NO:126, SEQ ID NO:127, SEQ ID NO:128, SEQ ID NO:129, SEQ ID NO:130, SEQ ID

NO:131, SEQ ID NO:132, SEQ ID NO:133, SEQ ID NO:134 (SEQ ID NO:1-134), and fragments thereof.

The invention further provides a substantially purified variant having at least 90% amino acid identity to at least one of the amino acid sequences selected from the group consisting of SEQ ID NO:1-134, and fragments thereof. The invention also provides an isolated and purified polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-134, and fragments thereof. The invention also includes an isolated and purified polynucleotide variant having at least 90% polynucleotide sequence identity to the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-134, and fragments thereof.

Additionally, the invention provides an isolated and purified polynucleotide which hybridizes under stringent conditions to the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-134, and fragments thereof. The invention also provides an isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide encoding the polypeptide comprising the amino acid sequence selected from the group consisting of SEQ ID NO:1-134, and fragments thereof.

The invention also provides an isolated and purified polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:135, SEQ ID NO:136, SEQ ID NO:137, SEQ ID NO:138, SEQ ID NO:139, SEQ ID NO:140, SEQ ID NO:141, SEQ ID NO:142, SEQ ID NO:143, SEQ ID NO:144, SEQ ID NO:145, SEQ ID NO:146, SEQ ID NO:147, SEQ ID NO:148, SEQ ID NO:149, SEQ ID NO:150, SEQ ID NO:151, SEQ ID NO:152, SEQ ID NO:153, SEQ ID NO:154, SEQ ID NO:155, SEQ ID NO:156, SEQ ID NO:157, SEQ ID NO:158, SEQ ID NO:159, SEQ ID NO:160, SEQ ID NO:161, SEQ ID NO:162, SEQ ID NO:163, SEQ ID NO:164, SEQ ID NO:165, SEQ ID NO:166, SEQ ID NO:167, SEQ ID NO:168, SEQ ID NO:169, SEQ ID NO:170, SEQ ID NO:171, SEQ ID NO:172, SEQ ID NO:173, SEQ ID NO:174, SEQ ID NO:175, SEQ ID NO:176, SEQ ID NO:177, SEQ ID NO:178, SEQ ID NO:179, SEQ ID NO:180, SEQ ID NO:181, SEQ ID NO:182, SEQ ID NO:183, SEQ ID NO:184, SEQ ID NO:185, SEQ ID NO:186, SEQ ID NO:187, SEQ ID NO:188, SEQ ID NO:189, SEQ ID NO:190, SEQ ID NO:191, SEQ ID NO:192, SEQ ID NO:193, SEQ ID NO:194, SEQ ID NO:195, SEQ ID

NO:196, SEQ ID NO:197, SEQ ID NO:198, SEQ ID NO:199, SEQ ID NO:200, SEQ ID NO:201, SEQ ID NO:202, SEQ ID NO:203, SEQ ID NO:204, SEQ ID NO:205, SEQ ID NO:206, SEQ ID NO:207, SEQ ID NO:208, SEQ ID NO:209, SEQ ID NO:210, SEQ ID NO:211, SEQ ID NO:212, SEQ ID NO:213, SEQ ID NO:214, SEQ ID NO:215, SEQ ID NO:216, SEQ ID NO:217, SEQ ID NO:218, SEQ ID NO:219, SEQ ID NO:220, SEQ ID NO:221, SEQ ID NO:222, SEQ ID NO:223, SEQ ID NO:224, SEQ ID NO:225, SEQ ID NO:226, SEQ ID NO:227, SEQ ID NO:228, SEQ ID NO:229, SEQ ID NO:230, SEQ ID NO:231, SEQ ID NO:232, SEQ ID NO:233, SEQ ID NO:234, SEQ ID NO:235, SEQ ID NO:236, SEQ ID NO:237, SEQ ID NO:238, SEQ ID NO:239, SEQ ID NO:240, SEQ ID NO:241, SEQ ID NO:242, SEQ ID NO:243, SEQ ID NO:244, SEQ ID NO:245, SEQ ID NO:246, SEQ ID NO:247, SEQ ID NO:248, SEQ ID NO:249, SEQ ID NO:250, SEQ ID NO:251, SEQ ID NO:252, SEQ ID NO:253, SEQ ID NO:254, SEQ ID NO:255, SEQ ID NO:256, SEQ ID NO:257, SEQ ID NO:258, SEQ ID NO:259, SEQ ID NO:260, SEQ ID NO:261, SEQ ID NO:262, SEQ ID NO:263, SEQ ID NO:264, SEQ ID NO:265, SEQ ID NO:266, SEQ ID NO:267, SEQ ID NO:268 (SEQ ID NO:135-268), and fragments thereof. The invention further provides an isolated and purified polynucleotide variant having at least 90% polynucleotide sequence identity to the polynucleotide sequence selected from the group consisting of SEQ ID NO:135-268, and fragments thereof. The invention also provides an isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:135-268, and fragments thereof.

The invention also provides a method for detecting a polynucleotide in a sample containing nucleic acids, the method comprising the steps of (a) hybridizing the complement of the polynucleotide sequence to at least one of the polynucleotides of the sample, thereby forming a hybridization complex; and (b) detecting the hybridization complex, wherein the presence of the hybridization complex correlates with the presence of a polynucleotide in the sample. In one aspect, the method further comprises amplifying the polynucleotide prior to hybridization.

The invention further provides an expression vector containing at least a fragment of the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-134, and fragments thereof. In another aspect, the expression vector is contained within a host cell.



The invention also provides a method for producing a polypeptide, the method comprising the steps of: (a) culturing the host cell containing an expression vector containing at least a fragment of a polynucleotide under conditions suitable for the expression of the polypeptide; and (b) recovering the polypeptide from the host cell  
5 culture.

The invention also provides a pharmaceutical composition comprising a substantially purified polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO:1-134, and fragments thereof, in conjunction with a suitable pharmaceutical carrier.

10 The invention further includes a purified antibody which binds to a polypeptide selected from the group consisting of SEQ ID NO:1-134, and fragments thereof. The invention also provides a purified agonist and a purified antagonist to the polypeptide.

The invention also provides a method for treating or preventing a disorder associated with decreased expression or activity of HSPP, the method comprising  
15 administering to a subject in need of such treatment an effective amount of a pharmaceutical composition comprising a substantially purified polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO:1-134, and fragments thereof, in conjunction with a suitable pharmaceutical carrier.

The invention also provides a method for treating or preventing a disorder  
20 associated with increased expression or activity of HSPP, the method comprising administering to a subject in need of such treatment an effective amount of an antagonist of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-134, and fragments thereof.

25

#### BRIEF DESCRIPTION OF THE TABLE

Table 1 shows nucleotide and polypeptide sequence identification numbers (SEQ ID NO), clone identification numbers (clone ID), cDNA libraries, and cDNA fragments used to assemble full-length sequences encoding HSPP.

30 Table 2 shows features of each polypeptide sequence, including predicted signal peptide sequences, and methods and algorithms used for identification of HSPP.

Synthesizer (Perkin-Elmer). Various fragments of HSPP may be synthesized separately and then combined to produce the full length molecule.

## THERAPEUTICS

5 Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of HSPP and signal peptide sequences. In addition, chemical and structural similarity, in the context of sequences and motifs, exists between HSPP-66 and prostatic steriod-binding C3 precursor from rat (GI 206453); between HSPP-68 and TWIK-related acid-sensitive K<sup>+</sup>channel from human (GI 2465542); and between HSPP-92  
10 and tyrosine specific protein phosphatases (PROSITE PDOC00323). In addition, the expression of HSPP is closely associated with proliferative, cancerous, inflamed, cardiovascular, nervous, reproductive, hematopoietic/immune, and developmental tissue. Therefore, HSPP appears to play a role in cell proliferative disorders including cancer; inflammation; and cardiovascular,  
15 neurological, reproductive, and developmental disorders. In the treatment of cell proliferative disorders including cancer; inflammation; and cardiovascular, neurological, reproductive, and developmental disorders associated with increased HSPP expression or activity, it is desirable to decrease the expression or activity of HSPP. In the treatment of the above conditions associated with decreased HSPP expression or activity, it is desirable  
20 to increase the expression or activity of HSPP.

Therefore, in one embodiment, HSPP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HSPP. Examples of such disorders include, but are not limited to, cell proliferative disorders such as actinic keratosis, arteriosclerosis, atherosclerosis,  
25 bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia,  
30 gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; inflammatory disorders, such as acquired immunodeficiency syndrome (AIDS), Addison's

- disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis,
- 5 dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis,
- 10 polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; cardiovascular disorders including
- 15 disorders of the blood vessels such as arteriovenous fistula, atherosclerosis, hypertension, vasculitis, Raynaud's disease, aneurysms, arterial dissections, varicose veins, thrombophlebitis and phlebothrombosis, and vascular tumors; disorders of the heart such as congestive heart failure, ischemic heart disease, angina pectoris, myocardial infarction, hypertensive heart disease, degenerative valvular heart disease, calcific aortic valve
- 20 stenosis, congenitally bicuspid aortic valve, mitral annular calcification, mitral valve prolapse, rheumatic fever and rheumatic heart disease, infective endocarditis, nonbacterial thrombotic endocarditis, endocarditis of systemic lupus erythematosus, carcinoid heart disease, cardiomyopathy, myocarditis, pericarditis, neoplastic heart disease, and congenital heart disease; and disorders of the lungs such as congenital lung anomalies, atelectasis,
- 25 pulmonary congestion and edema, pulmonary embolism, pulmonary hemorrhage, pulmonary infarction, pulmonary hypertension, vascular sclerosis, obstructive pulmonary disease, restrictive pulmonary disease, chronic obstructive pulmonary disease, emphysema, chronic bronchitis, bronchial asthma, bronchiectasis, bacterial pneumonia, viral and mycoplasmal pneumonia, lung abscess, pulmonary tuberculosis, diffuse
- 30 interstitial diseases, pneumoconioses, sarcoidosis, idiopathic pulmonary fibrosis, desquamative interstitial pneumonitis, hypersensitivity pneumonitis, pulmonary eosinophilia bronchiolitis obliterans-organizing pneumonia, diffuse pulmonary

hemorrhage syndromes, Goodpasture's syndromes, idiopathic pulmonary hemosiderosis, pulmonary involvement in collagen-vascular disorders, pulmonary alveolar proteinosis, lung tumors, inflammatory and noninflammatory pleural effusions, pneumothorax, and pleural tumors; neurological disorders such as epilepsy, ischemic cerebrovascular disease, 5 stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, 10 suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease; prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome; fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other 15 developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis; inherited, metabolic, endocrine, and toxic myopathies; myasthenia gravis, periodic paralysis; mental disorders including mood, 20 anxiety, and schizophrenic disorders; akathisia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, and Tourette's disorder; reproductive disorders such as disorders of prolactin production; infertility, including tubal disease, ovulatory defects, and endometriosis; disruptions of the estrous cycle, disruptions of the menstrual cycle, polycystic ovary syndrome, ovarian 25 hyperstimulation syndrome, endometrial and ovarian tumors, uterine fibroids, autoimmune disorders, ectopic pregnancies, and teratogenesis; cancer of the breast, fibrocystic breast disease, and galactorrhea; disruptions of spermatogenesis, abnormal sperm physiology, cancer of the testis, cancer of the prostate, benign prostatic hyperplasia, prostatitis, Peyronie's disease, carcinoma of the male breast, and gynecomastia; and developmental 30 disorders, such as renal tubular acidosis, anemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular dystrophy, epilepsy, gonadal dysgenesis, WAGR syndrome (Wilms' tumor, aniridia, genitourinary abnormalities, and mental

retardation), Smith-Magenis syndrome, myelodysplastic syndrome, hereditary mucoepithelial dysplasia, hereditary keratodermas, hereditary neuropathies such as Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism, hydrocephalus, seizure disorders such as Sydenham's chorea and cerebral palsy, spina bifida, anencephaly, craniorachischisis, congenital glaucoma, cataract, and sensorineural hearing loss.

In another embodiment, a vector capable of expressing HSPP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HSPP including, but not limited to, those described above.

In a further embodiment, a pharmaceutical composition comprising a substantially purified HSPP in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HSPP including, but not limited to, those provided above.

In still another embodiment, an agonist which modulates the activity of HSPP may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HSPP including, but not limited to, those listed above.

In a further embodiment, an antagonist of HSPP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of HSPP. Examples of such disorders include, but are not limited to, those described above. In one aspect, an antibody which specifically binds HSPP may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissue which express HSPP.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding HSPP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of HSPP including, but not limited to, those described above.

In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act

synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of HSPP may be produced using methods which are generally  
5 known in the art. In particular, purified HSPP may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind HSPP. Antibodies to HSPP may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression  
10 library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are especially preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with HSPP or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species,  
15 various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

20 It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to HSPP have an amino acid sequence consisting of at least about 5 amino acids, and, more preferably, of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein and contain the entire amino acid sequence of a small,  
25 naturally occurring molecule. Short stretches of HSPP amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to HSPP may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture.  
30 These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42;

Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. 80:2026-2030; and Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule  
5 with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; and Takeda, S. et al. (1985) Nature 314:452-454.)

Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce HSPP-specific single chain  
10 antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton D.R. (1991) Proc. Natl. Acad. Sci. 88:10134-10137.)

Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly  
15 specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. 86: 3833-3837; Winter, G. et al. (1991) Nature 349:293-299.)

Antibody fragments which contain specific binding sites for HSPP may also be generated. For example, such fragments include, but are not limited to, F(ab')<sub>2</sub> fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by  
20 reducing the disulfide bridges of the F(ab')<sub>2</sub> fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) Science 246:1275-1281.)

Various immunoassays may be used for screening to identify antibodies having the  
25 desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between HSPP and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering HSPP  
30 epitopes is preferred, but a competitive binding assay may also be employed (Pound, supra).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for HSPP. Affinity is expressed as an association constant,  $K_a$ , which is defined as the molar concentration of HSPP-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The  $K_a$  determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple HSPP epitopes, represents the average affinity, or avidity, of the antibodies for HSPP. The  $K_a$  determined for a preparation of monoclonal antibodies, which are monospecific for a particular HSPP epitope, represents a true measure of affinity. High-affinity antibody preparations with  $K_a$  ranging from about  $10^9$  to  $10^{12}$  L/mole are preferred for use in immunoassays in which the HSPP-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with  $K_a$  ranging from about  $10^6$  to  $10^7$  L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of HSPP, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington, DC; Liddell, J. E. and Cryer, A. (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is preferred for use in procedures requiring precipitation of HSPP-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available. (See, e.g., Catty, supra, and Coligan et al. supra.)

In another embodiment of the invention, the polynucleotides encoding HSPP, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, the complement of the polynucleotide encoding HSPP may be used in situations in which it would be desirable to block the transcription of the mRNA. In particular, cells may be transformed with sequences complementary to polynucleotides encoding HSPP. Thus, complementary molecules or fragments may be used to modulate HSPP activity, or to achieve regulation of gene function. Such technology is now well known in the art, and



sense or antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding HSPP.

Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide  
5 sequences to the targeted organ, tissue, or cell population. Methods which are well known to those skilled in the art can be used to construct vectors to express nucleic acid sequences complementary to the polynucleotides encoding HSPP. (See, e.g., Sambrook, supra; Ausubel, 1995, supra.)

Genes encoding HSPP can be turned off by transforming a cell or tissue with  
10 expression vectors which express high levels of a polynucleotide, or fragment thereof, encoding HSPP. Such constructs may be used to introduce untranslatable sense or antisense sequences into a cell. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until they are disabled by endogenous nucleases. Transient expression may last for a month or more with a non-replicating  
15 vector, and may last even longer if appropriate replication elements are part of the vector system.

As mentioned above, modifications of gene expression can be obtained by designing complementary sequences or antisense molecules (DNA, RNA, or PNA) to the control, 5', or regulatory regions of the gene encoding HSPP. Oligonucleotides derived  
20 from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, are preferred. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex  
25 DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

30 Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by

endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding HSPP.

Specific ribozyme cleavage sites within any potential RNA target are initially  
5 identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be  
10 evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase  
15 phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding HSPP. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell  
20 lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is  
25 inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

Many methods for introducing vectors into cells or tissues are available and  
30 equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or

by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) *Nature Biotechnology* 15:462-466.)

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as dogs, cats, cows, horses, rabbits, monkeys, and most preferably, humans.

An additional embodiment of the invention relates to the administration of a pharmaceutical or sterile composition, in conjunction with a pharmaceutically acceptable carrier, for any of the therapeutic effects discussed above. Such pharmaceutical compositions may consist of HSPP, antibodies to HSPP, and mimetics, agonists, antagonists, or inhibitors of HSPP. The compositions may be administered alone or in combination with at least one other agent, such as a stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical carrier including, but not limited to, saline, buffered saline, dextrose, and water. The compositions may be administered to a patient alone, or in combination with other agents, drugs, or hormones.

The pharmaceutical compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Further details on techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA).

Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

Pharmaceutical preparations for oral use can be obtained through combining active compounds with solid excipient and processing the resultant mixture of granules (optionally, after grinding) to obtain tablets or dragee cores. Suitable auxiliaries can be

added, if desired. Suitable excipients include carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, and sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; gums, including  
5 arabic and tragacanth; and proteins, such as gelatin and collagen. If desired, disintegrating or solubilizing agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, and alginic acid or a salt thereof, such as sodium alginate.

Dragee cores may be used in conjunction with suitable coatings, such as concentrated sugar solutions, which may also contain gum arabic, talc,  
10 polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, i.e., dosage.

Pharmaceutical preparations which can be used orally include push-fit capsules  
15 made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with fillers or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with  
20 or without stabilizers.

Pharmaceutical formulations suitable for parenteral administration may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such  
25 as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils, such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate, triglycerides, or liposomes. Non-lipid polycationic amino polymers may also be used for delivery. Optionally, the suspension may also contain  
30 suitable stabilizers or agents to increase the solubility of the compounds and allow for the preparation of highly concentrated solutions.

For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

The pharmaceutical compositions of the present invention may be manufactured in a manner that is known in the art, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes.

The pharmaceutical composition may be provided as a salt and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, and succinic acid. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preferred preparation may be a lyophilized powder which may contain any or all of the following: 1 mM to 50 mM histidine, 0.1% to 2% sucrose, and 2% to 7% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. For administration of HSPP, such labeling would include amount, frequency, and method of administration.

Pharmaceutical compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells or in animal models such as mice, rats, rabbits, dogs, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example HSPP or fragments thereof, antibodies of HSPP, and agonists, antagonists or inhibitors of HSPP, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED<sub>50</sub> (the dose therapeutically

effective in 50% of the population) or  $LD_{50}$  (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, and it can be expressed as the  $LD_{50}/ED_{50}$  ratio. Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the  $ED_{50}$  with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about 0.1  $\mu\text{g}$  to 100,000  $\mu\text{g}$ , up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

## DIAGNOSTICS

In another embodiment, antibodies which specifically bind HSPP may be used for the diagnosis of disorders characterized by expression of HSPP, or in assays to monitor patients being treated with HSPP or agonists, antagonists, or inhibitors of HSPP. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for HSPP include methods which utilize the antibody and a label to detect HSPP in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled

<210> 24  
 <211> 311  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte Clone No: 1831477

<400> 24  
 Met Gly Val Pro Thr Ala Pro Glu Ala Gly Ser Trp Arg Trp Gly  
 1 5 10 15  
 Ser Leu Leu Phe Ala Leu Phe Leu Ala Ala Ser Leu Gly Pro Val  
 20 25 30  
 Ala Ala Phe Lys Val Ala Thr Pro Tyr Ser Leu Tyr Val Cys Pro  
 35 40 45  
 Glu Gly Gln Asn Val Thr Leu Thr Cys Arg Leu Leu Gly Pro Val  
 50 55 60  
 Asp Lys Gly His Asp Val Thr Phe Tyr Lys Thr Trp Tyr Arg Ser  
 65 70 75  
 Ser Arg Gly Glu Val Gln Thr Cys Ser Glu Arg Arg Pro Ile Arg  
 80 85 90  
 Asn Leu Thr Phe Gln Asp Leu His Leu His His Gly Gly His Gln  
 95 100 105  
 Ala Ala Asn Thr Ser His Asp Leu Ala Gln Arg His Gly Leu Glu  
 110 115 120  
 Ser Ala Ser Asp His His Gly Asn Phe Ser Ile Thr Met Arg Asn  
 125 130 135  
 Leu Thr Leu Leu Asp Ser Gly Leu Tyr Cys Cys Leu Val Val Glu  
 140 145 150  
 Ile Arg His His His Ser Glu His Arg Val His Gly Ala Met Glu  
 155 160 165  
 Leu Gln Val Gln Thr Gly Lys Asp Ala Pro Ser Asn Cys Val Val  
 170 175 180  
 Tyr Pro Ser Ser Ser Gln Glu Ser Glu Asn Ile Thr Ala Ala Ala  
 185 190 195  
 Leu Ala Thr Gly Ala Cys Ile Val Gly Ile Leu Cys Leu Pro Leu  
 200 205 210  
 Ile Leu Leu Leu Val Tyr Lys Gln Arg Gln Ala Ala Ser Asn Arg  
 215 220 225  
 Arg Ala Gln Glu Leu Val Arg Met Asp Ser Asn Ile Gln Gly Ile  
 230 235 240  
 Glu Asn Pro Gly Phe Glu Ala Ser Pro Pro Ala Gln Gly Ile Pro  
 245 250 255  
 Glu Ala Lys Val Arg His Pro Leu Ser Tyr Val Ala Gln Arg Gln  
 260 265 270  
 Pro Ser Glu Ser Gly Arg His Leu Leu Ser Glu Pro Ser Thr Pro  
 275 280 285  
 Leu Ser Pro Pro Gly Pro Gly Asp Val Phe Phe Pro Ser Leu Asp  
 290 295 300  
 Pro Val Pro Asp Ser Pro Asn Phe Glu Val Ile  
 305 310

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte Clone No: 1831477

&lt;400&gt; 158

```
ggagcacggc gctggggccg cccgcagcgc tcaactcgctc gcactcagtc gcgggaggct 60
tccccgcgcc ggccgcgtcc cgcgcgctcc ccggcaccag aagttcctct gcgcgtccga 120
cggcgacatg ggcgccccca cggccccgga ggccggcagc tggcgctggg gatccctgct 180
cttcgctctc ttccctggctg cgtccctagg tccggtggca gccttcaagg tcgccacgcc 240
gtattccctg tatgtctgtc ccgaggggca gaacgtcacc ctcacctgca ggctcttggg 300
ccctgtggac aaagggcacg atgtgacctt ctacaagacg tggtaaccga gctcgagggg 360
cgaggtgcag acctgctcag agcgcgggcc catccgcaac ctacgttcc aggaacctca 420
cctgcaccat ggaggccacc aggctgccaa caccagccac gacctggctc agcgcacgg 480
gctggagtcg gcctccgacc accatggcaa cttctccatc accatgcgca acctgacct 540
gctggatagc ggccctctact gctgcctggt ggtggagatc aggcaccacc actcggagca 600
caggggtccat ggtgccatgg agctgcaggt gcagacaggc aaagatgcac catccaactg 660
tgtggtgtac ccattcctct cccaggagag tgaacacatc acggctgcag ccctggctac 720
gggtgcctgc atcgtaggaa tcctctgcct cccctcctc ctgctcctgg tctacaagca 780
aaggcaggca gcctccaacc gccgtgcccc ggagctgggt cggtatggaca gcaacattca 840
agggattgaa aaccccggtt ttgaagctc accacctgcc caggggatac ccgaggccaa 900
agtcaggcac cccctgtcct atgtggcccc gcggcagcct tctgagtctg ggccgacct 960
gctttcggag cccagcacc cctgtctcc tcaggcccc ggagacgtct tcttccatc 1020
cctggacctc gtccctgact ctccaaactt tgaggtcatc tagccagct gggggacagt 1080
gggctgttgt ggctgggtct ggggcagggt catttgagcc agggctggct ctgtgagtgg 1140
cctccttggc ctgcgccctg gttccctccc tctgtctctg ggctcagata ctgtgacatc 1200
ccagaagccc agccctcaa cccctctgga tgctacatgg ggatgctgga cggctcagcc 1260
cctgttccaa ggattttggg gtgctgagat tctcccctag agacctgaaa ttcaccagct 1320
acagatgcca aatgacttac atcttaagaa gtctcagaac gtccagccct tcagcagctc 1380
tcgttctgag acatgagcct tgggatgtgg cagcatcagt gggacaagat ggacactggg 1440
ccaccctccc aggcaccaga cacagggcac ggtggagaga cttctcccc gtggccgct 1500
tggtccccc gttttgcccg aggtgctct tctgtcagac ttctctttg taccacagt 1560
gctctggggc caggcctgcc tgcccaactg ccacgcgcac ctccccagc tgccctctac 1620
cagcagtttc tctgaagatc tgtcaacagg ttaagtcaat ctggggcttc cactgcctgc 1680
attccagtc ccagagcttg gtggtcccca aacgggaagt acatattggg gcatggtggc 1740
ctccgtgagc aatggtgtc ttgggcaatc tgaggccagg acagatgttg cccacccac 1800
tggagatggt gctgagggag gtgggtgggg ccttctggga aggtgagtgg agaggggac 1860
ctgcccccg cctccccat cccctactcc cactgctcag cgcgggcat tgcaagggtg 1920
ccacacaatg tcttgtccac cctgggacac ttctgagtat gaagcgggat gctattaaaa 1980
actacatggg gaaacaggtg caaaaaaaaa a 2011
```

&lt;210&gt; 159

&lt;211&gt; 480

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt;

&lt;222&gt; 440

&lt;223&gt; a or g or c or t, unknown, or other

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte Clone No: 1841607

&lt;400&gt; 159

```
cccacgcgtc cgaaaagaaa agaaaaaaga aaatggcctc atcttgtttc agcctcagtt 60
```